



Generating solutions of "impossible-tosolve" problems and simulating "impossible-to-simulate" models

Florent Krzakala
Espci-Paristech

L. Zdeborová (Los Alamos, LANL)





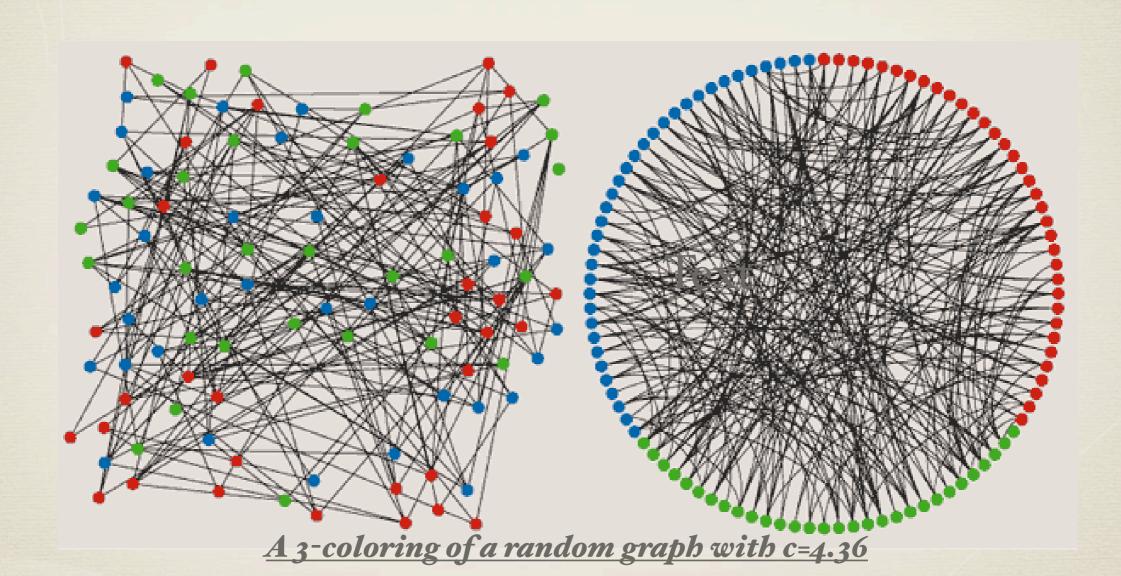
Planting!

Florent Krzakala Espci-Paristech

L. Zdeborová (Los Alamos, LANL)

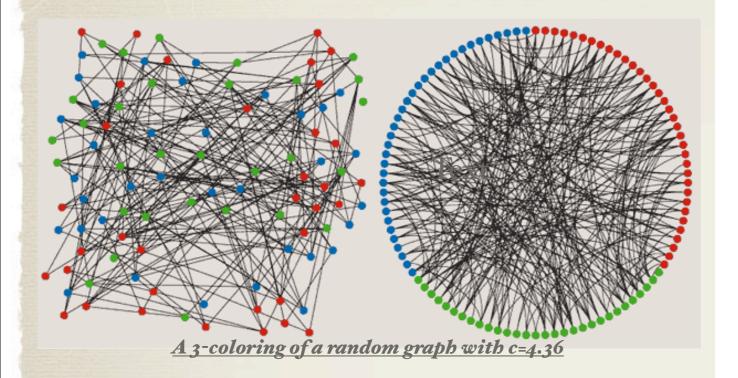
Some optimization problems such as COL and SAT are almost impossible to solve!

<u>ex:</u> Hard Instances of random graph coloring



Some optimization problems such as COL and SAT are almost impossible to solve!

<u>ex:</u> Hard Instances of random graph coloring

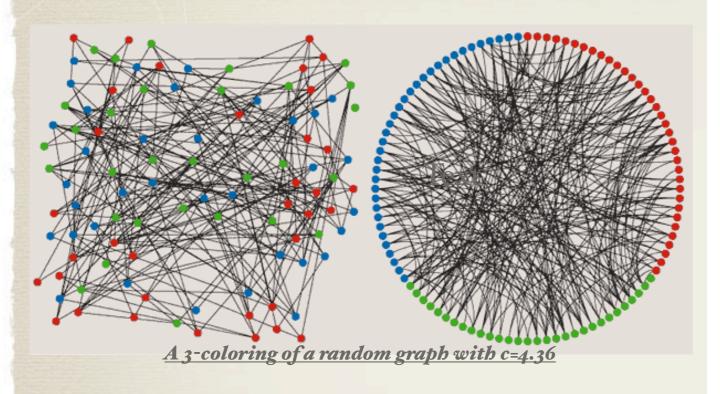


The " $q \log q$ " problem

D. Achlioptas et al. Nature 2005

Some optimization problems such as COL and SAT are almost impossible to solve!

<u>ex:</u> Hard Instances of random graph coloring



The " $q \log q$ " problem

- Consider q color (with q large enough) and a large random graph of average degree c
- W.h.p this graph is colorable if c<2q log q
- However, no algorithm is able to do so efficiently (polynomial) for c> q log q!

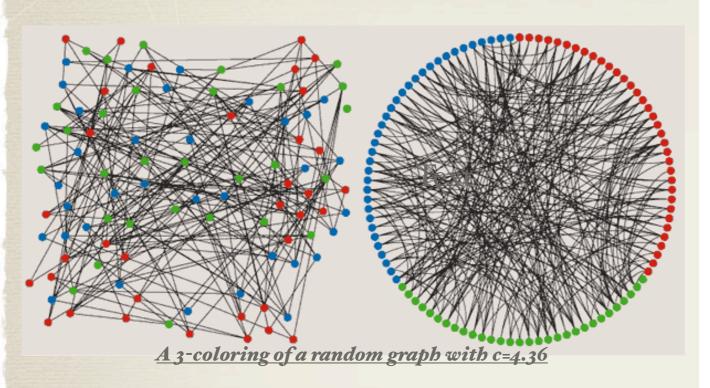
D. Achlioptas et al. Nature 2005

q log q

2q log q

Some optimization problems such as COL and SAT are almost impossible to solve!

<u>ex:</u> Hard Instances of random graph coloring



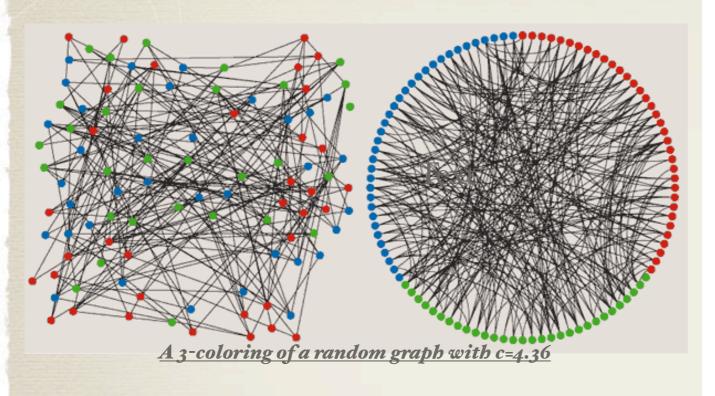
The " $q \log q$ " problem

- Consider q color (with q large enough) and a large random graph of average degree c
- W.h.p this graph is colorable if c<2q log q
- However, no algorithm is able to do so efficiently (polynomial) for c> q log q!

D. Achlioptas et al. Nature 2005

Some optimization problems such as COL and SAT are almost impossible to solve!

<u>ex:</u> Hard Instances of random graph coloring



The " $q \log q$ " problem

- Consider q color (with q large enough) and a large random graph of average degree c
- W.h.p this graph is colorable if c<2q log q
- However, no algorithm is able to do so efficiently (polynomial) for c> q log q!

Possible

D. Achlioptas et al. Nature 2005

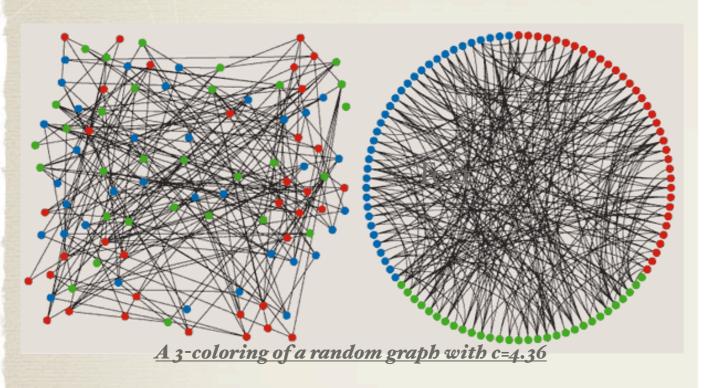
Average

degree c

Impossible

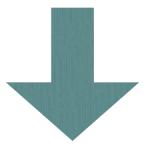
Some optimization problems such as COL and SAT are almost impossible to solve!

<u>ex:</u> Hard Instances of random graph coloring



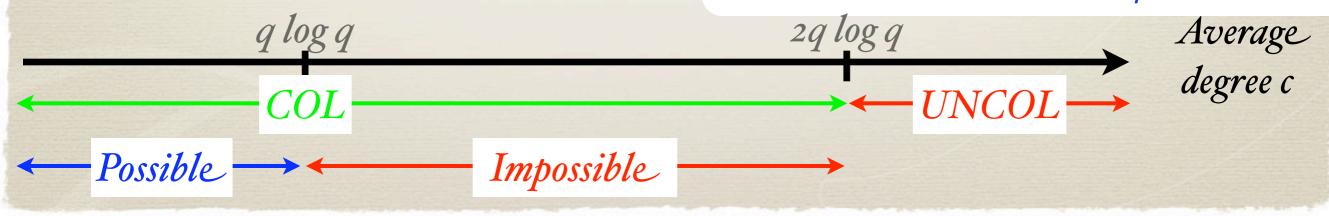
The " $q \log q$ " problem

- Consider q color (with q large enough) and a large random graph of average degree c
- W.h.p this graph is colorable if c<2q log q
- However, no algorithm is able to do so efficiently (polynomial) for c> q log q!



No one has ever seen the solution of, say 5-coloring, for large enough c and N=10⁶

D. Achlioptas et al. Nature 2005



Some optimization problems such as COL and SAT are also hard to sample

q log q

 $2q \log q$

Some optimization problems such as COL and SAT are also hard to sample

Consider the following "coloring" or "Potts-Antiferromagnet"

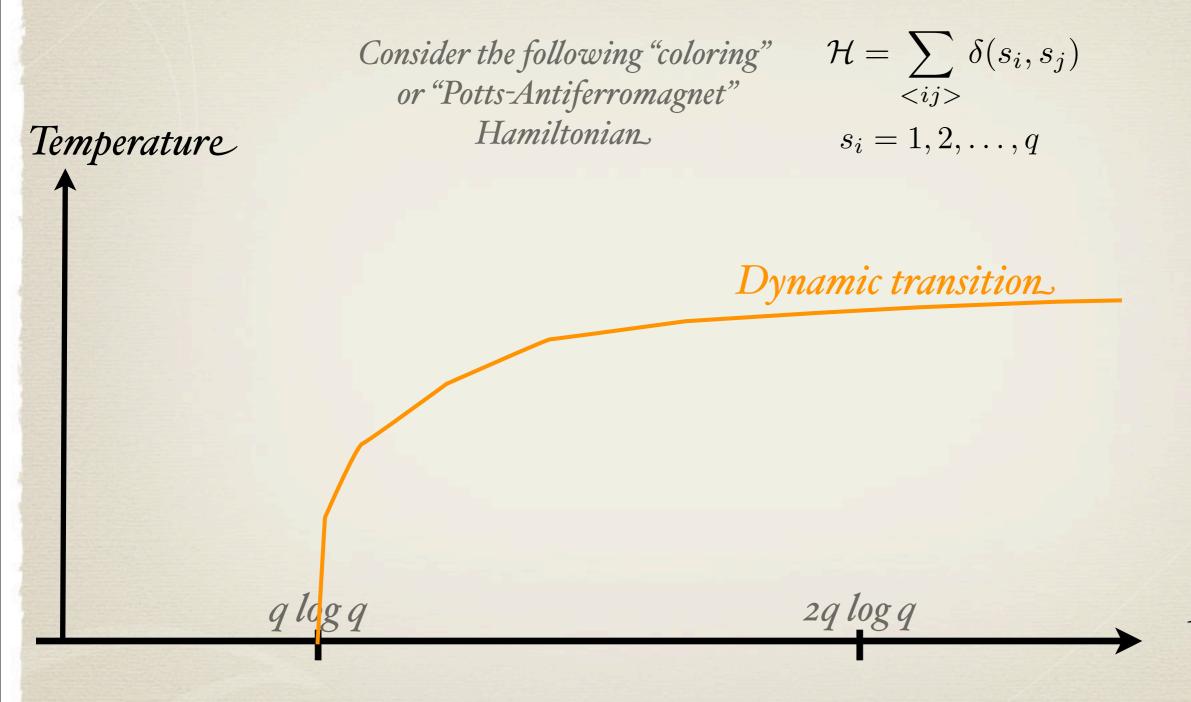
Hamiltonian_

$$\mathcal{H} = \sum_{\langle ij \rangle} \delta(s_i, s_j)$$
$$s_i = 1, 2, \dots, q$$

q log q

 $2q \log q$

Some optimization problems such as COL and SAT are also hard to sample



Some optimization problems such as COL and SAT are also hard to sample

Consider the following "coloring" or "Potts-Antiferromagnet"

Hamiltonian

$$\mathcal{H} = \sum_{\langle ij \rangle} \delta(s_i, s_j)$$
$$s_i = 1, 2, \dots, q$$

Possible-to-sample region_

Temperature

Dynamic transition_

Impossible-to-sample region_

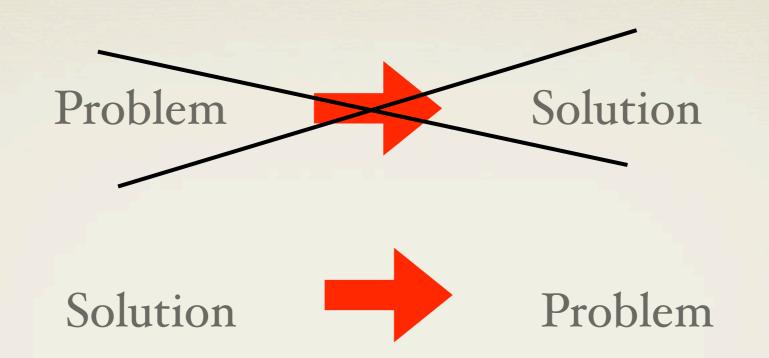
 $q \log q$

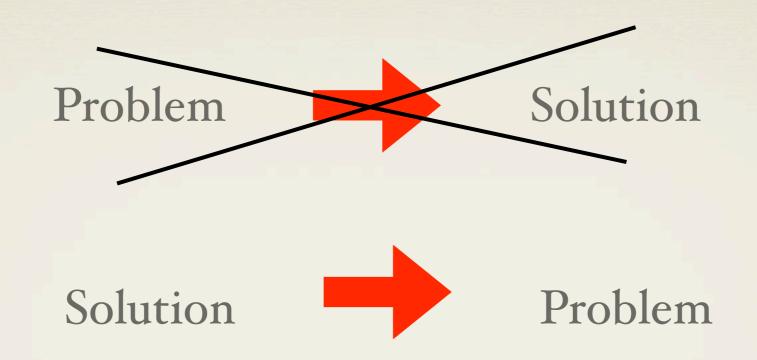
 $2q \log q$

Frustrating Intractable Problems

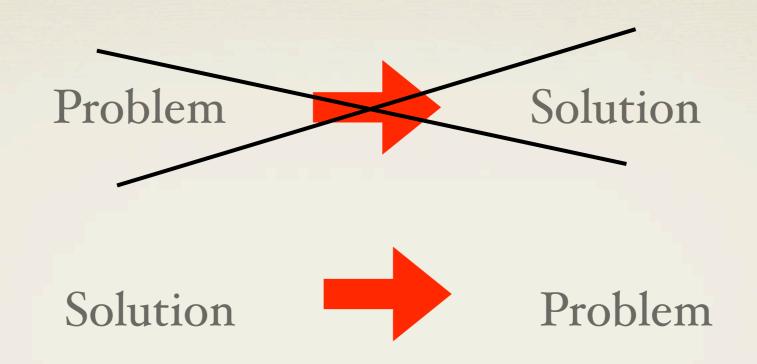
- * We know that some random problems DO have solutions, but we cannot find them!
- * Sampling and performing Monte-Carlo is even Harder!
- * Many predictions from statistical physics in random problems.... but impossible to test most of them!







Instead of choosing a problem, and looking for a solution....



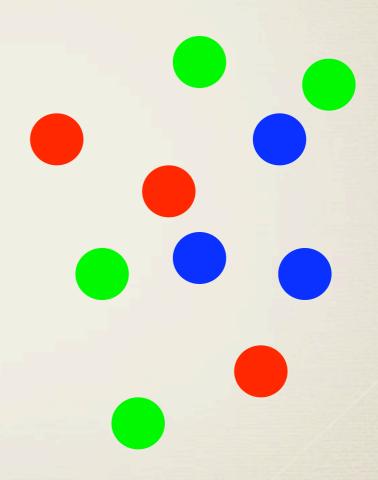
Instead of choosing a problem, and looking for a solution....



We choose a configuration/assignment and and look for a problem for which this is a solution!

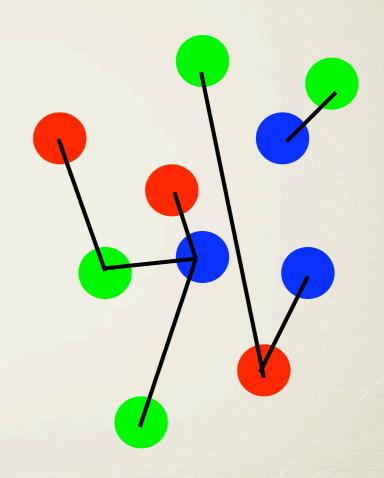
Consider the 3-coloring problem with N nodes and M links.

1) Color randomly the N nodes



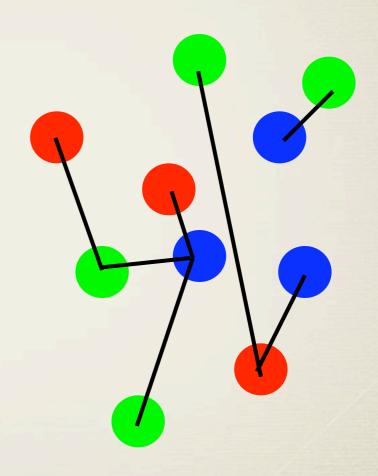
Consider the 3-coloring problem with N nodes and M links.

- 1) Color randomly the N nodes
- 11) Put the M links randomly such that the planted configuration is a proper coloring



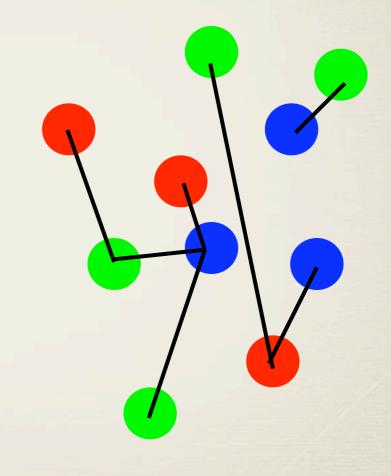
Consider the 3-coloring problem with N nodes and M links.

- 1) Color randomly the N nodes
- 11) Put the M links randomly such that the planted configuration is a proper coloring
- 111) Now, we have created a problem for which **we know** the solution



Consider the 3-coloring problem with N nodes and M links.

- 1) Color randomly the N nodes
- 11) Put the M links randomly such that the planted configuration is a proper coloring
- 111) Now, we have created a problem for which **we know** the solution



IV) We could also have prepared a configuration with a known cost/energy

Random ensemble

Choose a random graph with N nodes and M links

Planted ensemble

Choose a random coloring of N nodes

Choose a random graph such that this is a correct coloring...

Random ensemble

Choose a random graph with N nodes and M links

Planted ensemble

Choose a random coloring of N nodes

Choose a random graph such that this is a correct coloring...

* Is it really the same to look for a solution a random problem and to look for a random problem that matches a random solution?

Random ensemble

Choose a random graph with N nodes and M links

Planted ensemble

Choose a random coloring of N nodes

Choose a random graph such that this is a correct coloring...

- * Is it really the same to look for a solution a random problem and to look for a random problem that matches a random solution?
- * The surprising answer is: in some cases YES!

Random ensemble

Choose a random graph with N nodes and M links

Planted ensemble

Choose a random coloring of N nodes

Choose a random graph such that this is a correct coloring...

- * Is it really the same to look for a solution a random problem and to look for a random problem that matches a random solution?
- * The surprising answer is: in some cases YES!



Montanari and Semerjian, Jstat. '06 & Achlioptas and Coja-Oghlan, arXiv:0803.2122:

The two ensembles are asymptotically $(N\rightarrow \infty)$ equivalent for low enough degree c!

Random ensemble

Choose a random graph with N nodes and M links

Planted ensemble

Choose a random coloring of N nodes

Choose a random graph such that this is a correct coloring...

<u>Definition</u>: Two ensembles of random graphs are asymptotically equivalent if and only if in the thermodynamic limit every property which is almost surely true on a graph from one ensemble is also almost surely true on a graph from the other ensemble.

Some open questions:

- * Until which connectivity/degree c the planted and random ensembles are equivalent?
- * Is the planted ensemble interesting beyond this connectivity?
- * Can we generalize this approach to finite energy (coloring with a finite fraction of mistakes?)
- * How can we use a planted configuration?
- * What are the models where a "quiet" planting is possible?

In this talk:

1) A (brief) summary of a theory of "quiet" planting in random models

2) Using planted configurations for fast simulations.

The Planted Ensemble

The Planted Ensemble

* We use the formalism described in Zdeborová's talk

Main result

Consider a model where the annealed computation is correct in some region (high temperature or low degree)

$$f = -\frac{1}{N}\beta[\log Z]_{dis}$$



$$f = -\frac{1}{N}\beta[\log Z]_{dis}$$

$$f_{annealed} = -\frac{1}{N}\beta\log[Z]_{dis}$$

Main result

Consider a model where the annealed computation is correct in some region (high temperature or low degree)

$$f = -\frac{1}{N}\beta[\log Z]_{dis}$$

$$f_{annealed} = -\frac{1}{N}\beta\log[Z]_{dis}$$

Consider a model where a <u>factorized</u> (i.e. identical for all nodes)

Belief Propagation solution is correct in some region

(high temperature or low degree)

A list of models with "Quiet" planting!

This condition is fulfilled (at least in some region) for many models:

- Random q-coloring problem
- Random XOR-SAT
- Mean field spin glasses (e.g. Vianna-Bray, Sherrington-Kirkpatrick)
- Random 2-in-4 Sat
- Random Vertex-Cover (independent set)
- Any non disordered model on a random regular graphs

•....

A list of models with "Quiet" planting!

This condition is fulfilled (at least in some region) for many models:

- Random q-coloring problem
- Random XOR-SAT
- Mean field spin glasses (e.g. Vianna-Bray, Sherrington-Kirkpatrick)
- Random 2-in-4 Sat
- Random Vertex-Cover (independent set)
- Any non disordered model on a random regular graphs

•....

This condition is not fulfilled for:

•Random K-SAT

Main result

Consider a model where the annealed computation is correct in some region (high temperature or low degree)

Consider a model where a <u>factorized</u> (i.e. identical for all nodes)

Belief Propagation solution is correct in some region

(high temperature or low degree)

Main result

Consider a model where the annealed computation is correct in some region (high temperature or low degree)

Consider a model where a <u>factorized</u> (i.e. identical for all nodes)

Belief Propagation solution is correct in some region

(high temperature or low degree)



In the region where the two free energies are equal, the two ensembles are equivalent

Main result

Consider a model where the annealed computation is correct in some region (high temperature or low degree)

Consider a model where a <u>factorized</u> (i.e. identical for all nodes)

Belief Propagation solution is correct in some region

(high temperature or low degree)



In the region where the two free energies are equal, the two ensembles are equivalent



In the region where the two free energies are different, the planted configuration induces an additional "Gibbs" state (or BP fixed point)

A list of models with "Quiet" planting!

This condition is fulfilled (at least in some region) for many models:

- Random q-coloring problem
- Random XOR-SAT
- Mean field spin glasses (e.g. Vianna-Bray, Sherrington-Kirkpatrick)
- Random 2-in-4 Sat
- Random Vertex-Cover (independent set)
- Any non disordered model on a random regular graphs

•....

A list of models with "Quiet" planting!

This condition is fulfilled (at least in some region) for many models:

- Random q-coloring problem
- Random XOR-SAT
- Mean field spin glasses (e.g. Vianna-Bray, Sherrington-Kirkpatrick)
- Random 2-in-4 Sat
- Random Vertex-Cover (independent set)
- Any non disordered model on a random regular graphs

•....

For all these models, the cavity method allows to compute the value of the threshold beyond which f ≠ fannealed

"Phase transition"

A list of models with "Quiet" planting!

This condition is fulfilled (at least in some region) for many models:

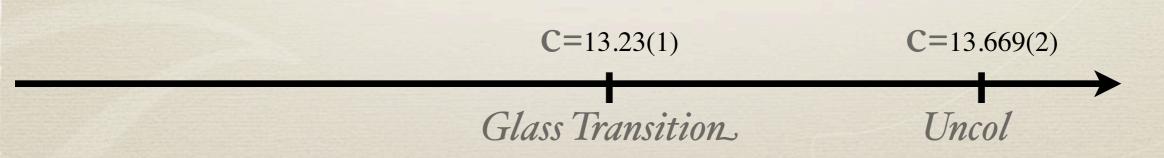
- Random q-coloring problem
- Random XOR-SAT
- Mean field spin glasses (e.g. Vianna-Bray, Sherrington-Kirkpatrick) In some models, the
- Random 2-in-4 Sat

- equivalence can be proven.
- Random Vertex-Cover (independ) Any non disordered model on a random regular graphs

For all these models, the cavity method allows to compute the value of the threshold beyond which f ≠ fannealed "Phase transition"

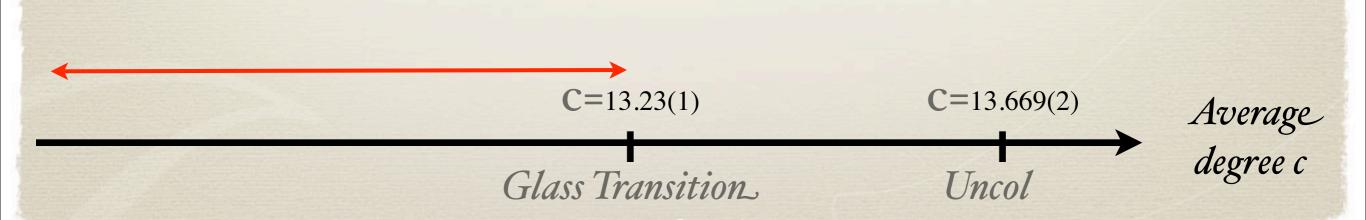
* Conjecture 1: the planted model is equivalent to the original one up to the point where the annealed solution is correct (for physicists: up to the static spin glass transition...) and the planted configuration is a "typical" one.

ex: 5-coloring of Erdös-Renyi random graphs



Average degree c

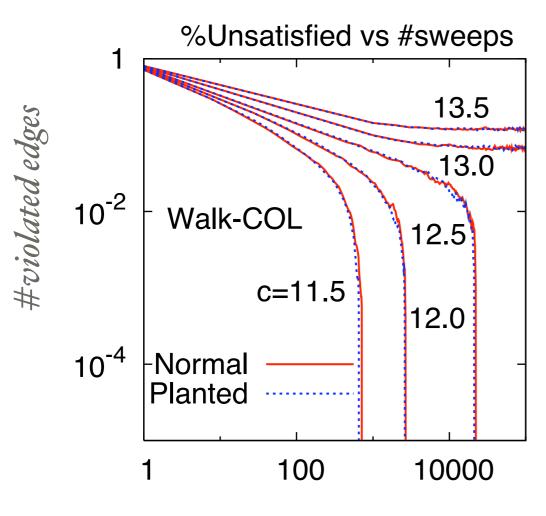
ex: 5-coloring of Erdös-Renyi random graphs



ex: 5-coloring of Erdös-Renyi random graphs

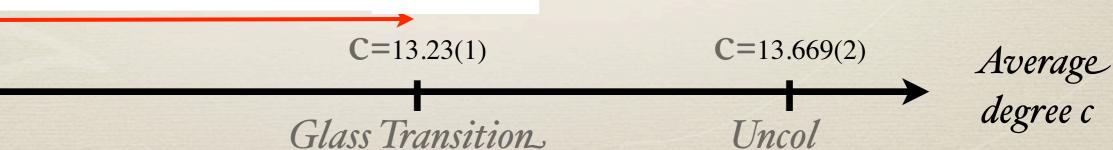
5-coloring using walkcol with

$$N=10^6$$

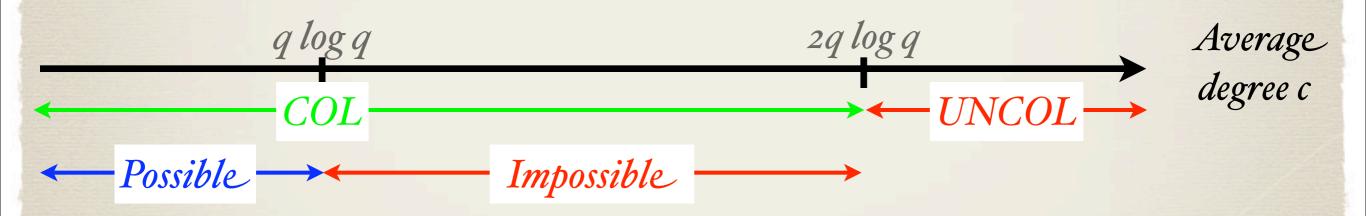


Number of flips /N

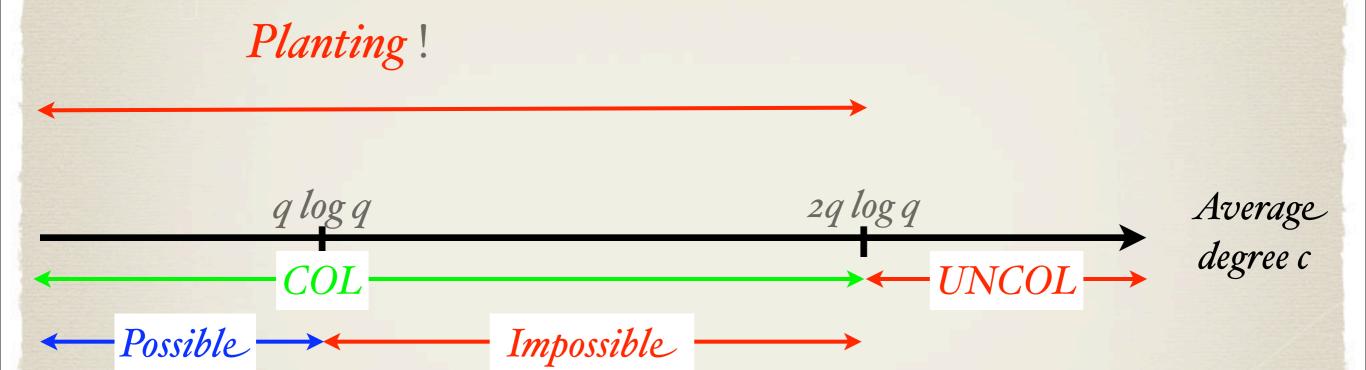
One can create impossible to solve problems of any size where the solution is known only by the creator



ex: q-coloring of Erdös-Renyi random graphs for large q



ex: q-coloring of Erdös-Renyi random graphs for large q



* Conjecture 1: the planted model is equivalent to the original one up to the point where the annealed solution is correct (for physicists: up to the spin glass transition...) and the planted configuration is a "typical" one

* Conjecture 1: the planted model is equivalent to the original one up to the point where the annealed solution is correct (for physicists: up to the spin glass transition...) and the planted configuration is a "typical" one

U. Feige, E. Mossel and D. Vilenchik.

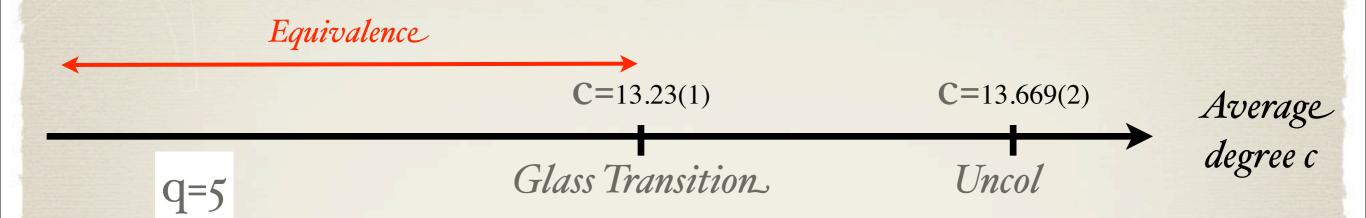
Proceedings of Random'06, LNCS 4410,



Planted configuration easy to find for large enough c

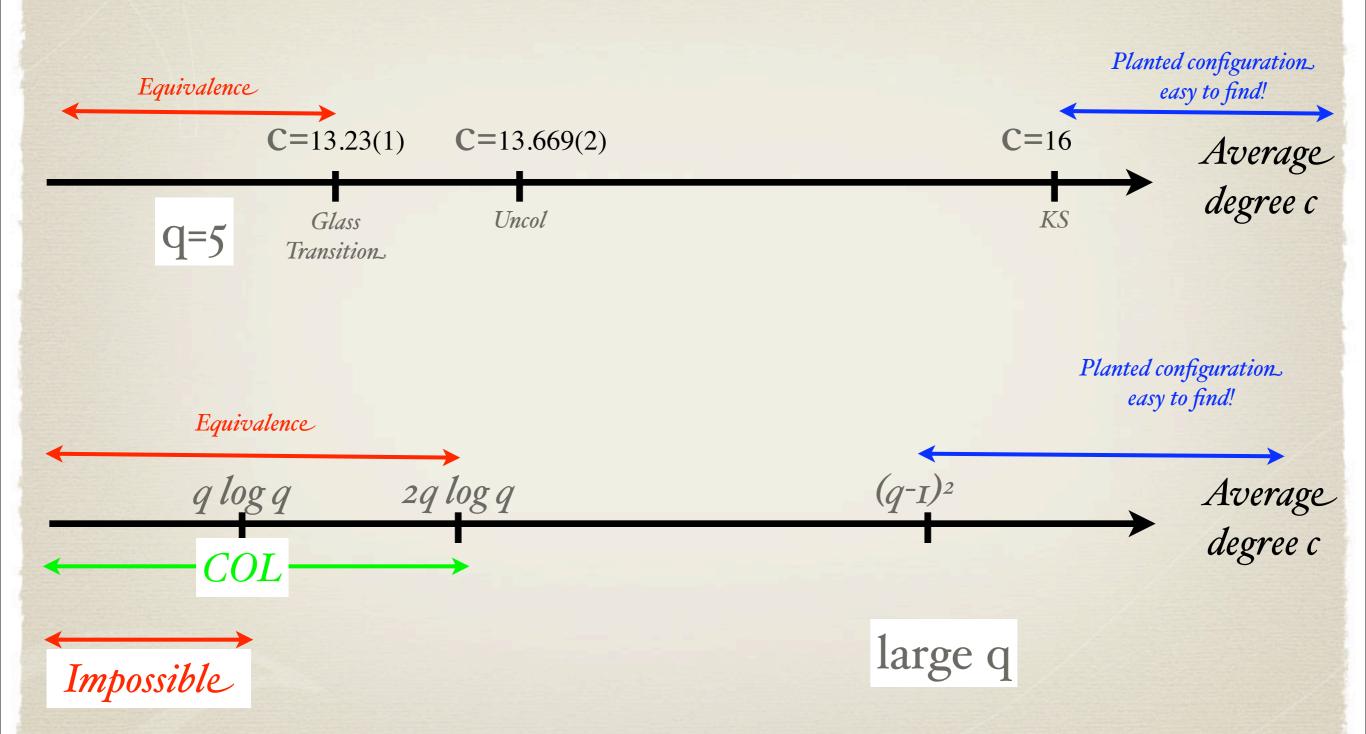
* Conjecture 1: the planted model is equivalent to the original one up to the point where the annealed solution is correct (for physicists: up to the spin glass transition...) and the planted configuration is a "typical" one

* Conjecture 2: Planted configuration are hard to find until the so-called Kesten-Stigum threshold, (for physicists: this is the <u>local spin glass</u> instability) beyond which they can be solved easily using BP.







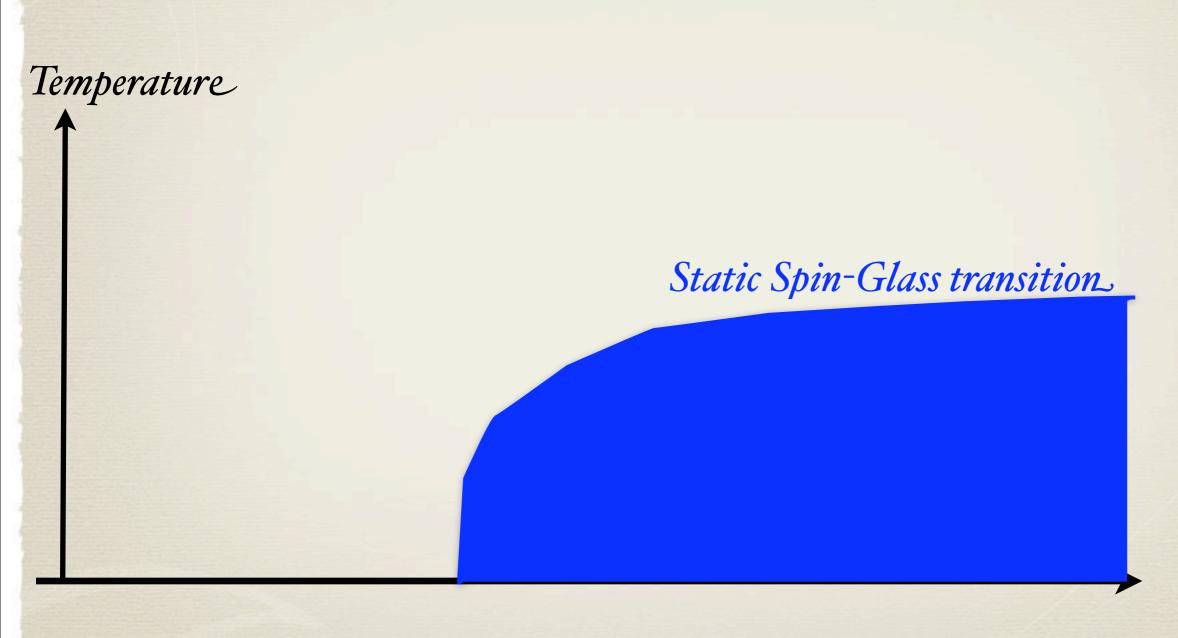


Simulating "impossible-to-simulate" models

How to perform simulations that are usually considered to be impossible?

Impossible-to-simulate problems

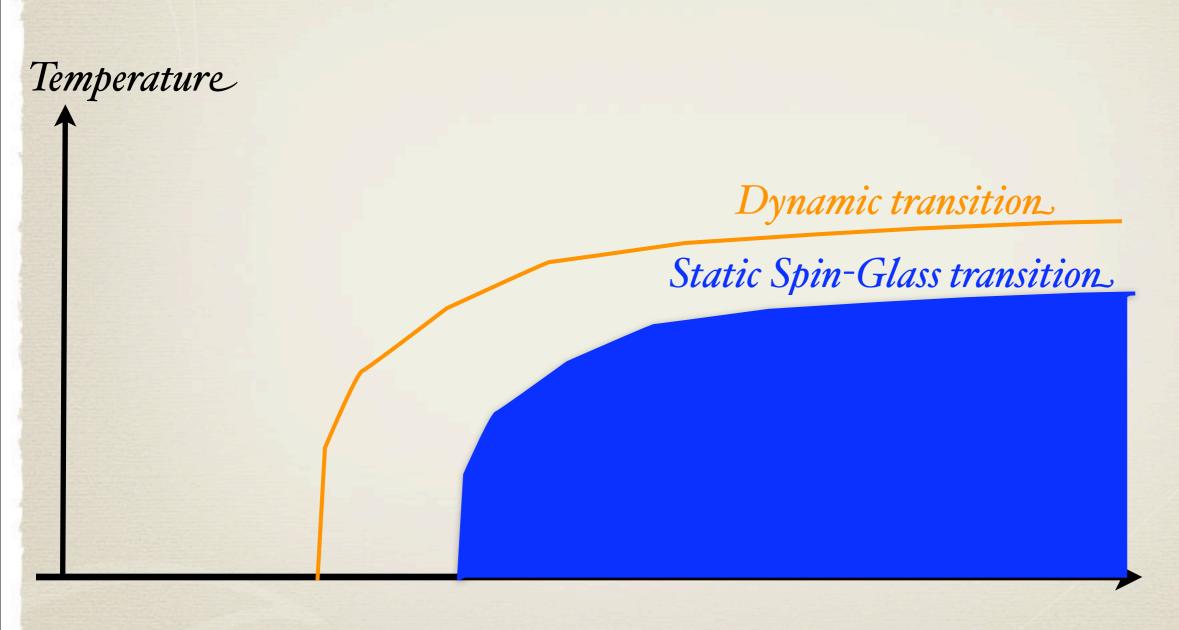
Random optimization problems & mean-field spin glasses



Average degree c

Impossible-to-simulate problems

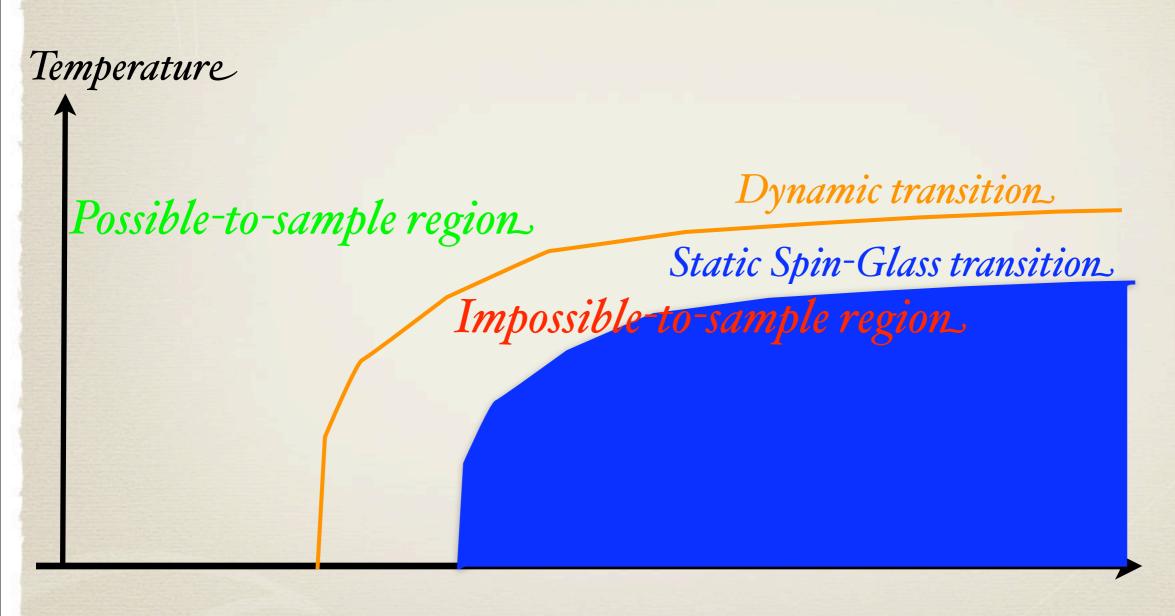
Random optimization problems & mean-field spin glasses



Average degree c

Impossible-to-simulate problems

Random optimization problems & mean-field spin glasses



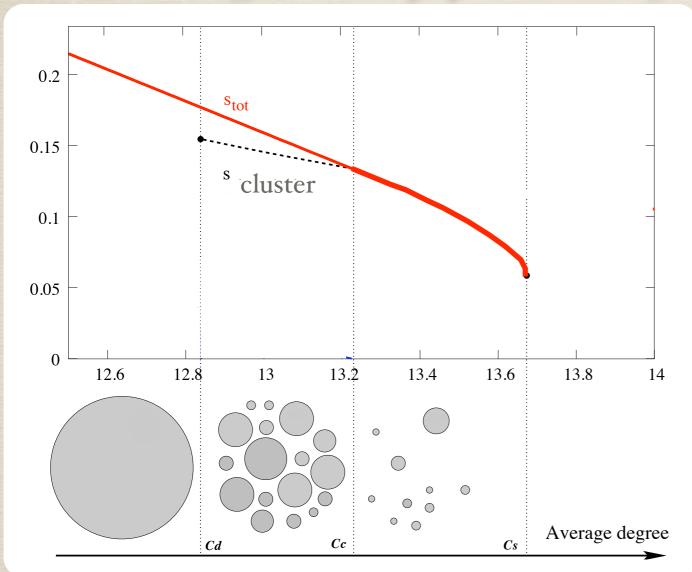
Average degree c

Modus operandi

- 1. Plant a configuration, create the graph such that the configuration satisfies all constraints
- 2. We now have a random instance and a "typical" equilibrium solution at zero temperature
- 3. We use it!

Example 1:

Testing the cavity predictions for the clustering transition



$$\psi_{factorized} = (\frac{1}{q}, \frac{1}{q} \dots \frac{1}{q})$$

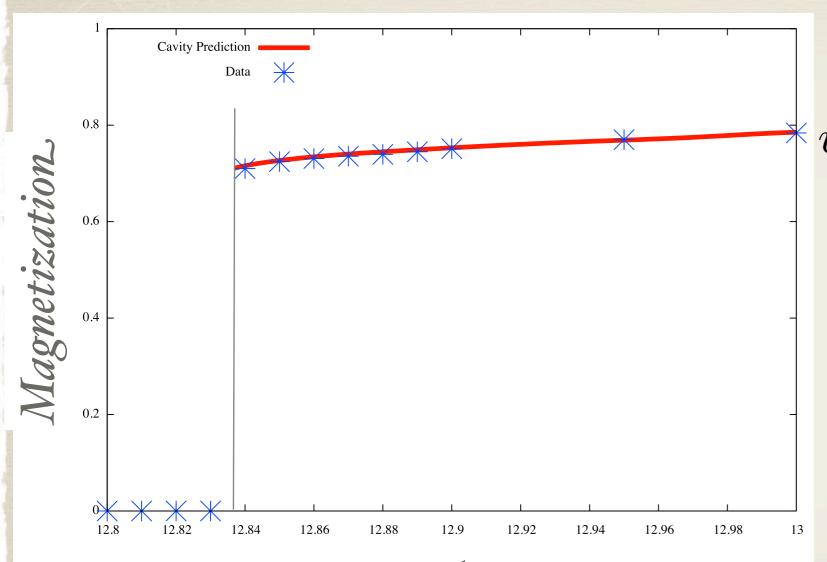
$$M = \frac{1}{qN} \sum_{graph} \sum_{c=1}^{q} \frac{\psi_{BP}^{c,i} - \frac{1}{q}}{1 - \frac{1}{q}}$$

FK, Montanari, Semerjian, Ricci-Tersenghi, Zdeborova, PNAS 07 & FK and Zdeborova, PRE 07

Prediction: beyond the so-called "dynamic" threshold, a non-trivial non-factorized fixed point of BP is obtained if one starts from an equilibrium configuration

Example 1:

Testing the cavity predictions for the clustering transition



$$\psi_{factorized} = (\frac{1}{q}, \frac{1}{q}, \dots, \frac{1}{q})$$

$$M = \frac{1}{qN} \sum_{graph} \sum_{c=1}^{q} \frac{\psi_{BP}^{c,i} - \frac{1}{q}}{1 - \frac{1}{q}}$$

Simulation with N=106

average degree

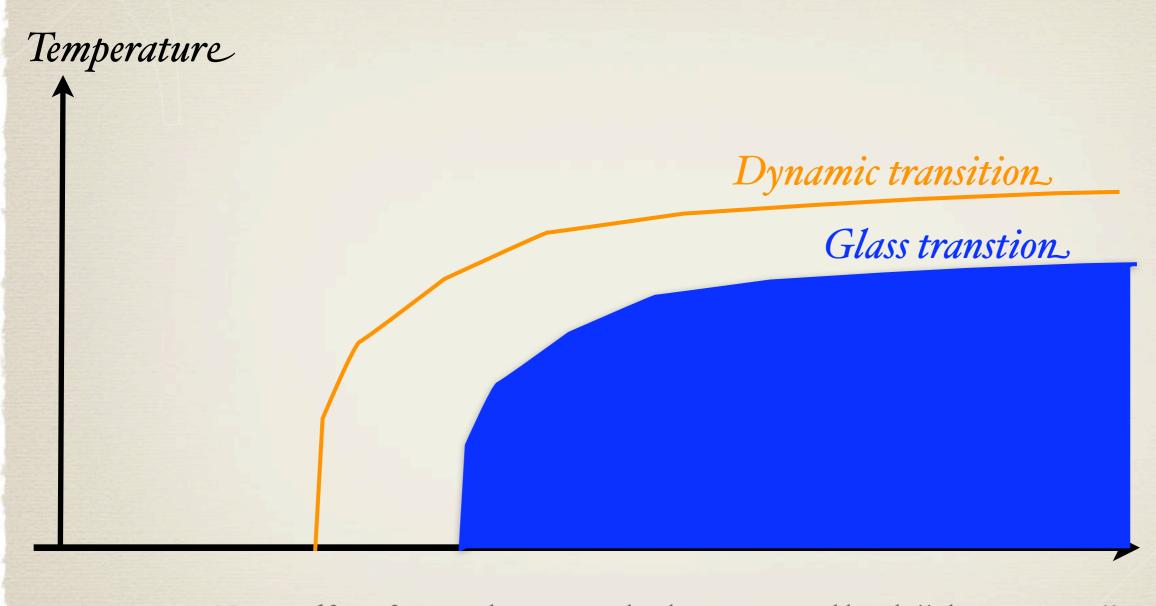
Prediction: beyond the so-called "dynamic" threshold, a

non-trivial non-factorized fixed point of BP is obtained if one starts from an equilibrium configuration

Modus operandi for finite temperature simulations

- 1. Plant a configuration, create the graph such that the configuration has exactly the equilibrium energy
- 2. We now have a random instance and a "typical" equilibrium solution at temperature T
- 3. We use it!

Testing the cavity predictions for the clustering transition



Average degree c

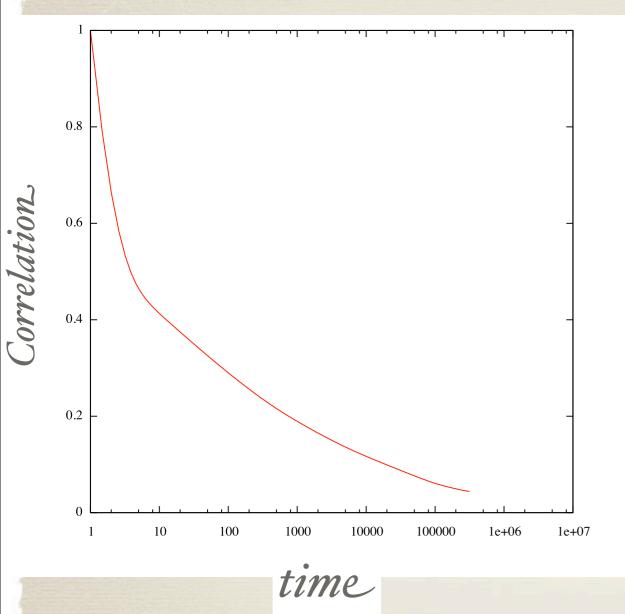
Testing the cavity predictions for the clustering transition

Usual Approach:

- 1) Start with a random initial condition.
- 2) compute the correlation function:

$$C(t) = \frac{1}{N} \sum_{i=1}^{N} S_i(t_{init} = 0) S_i(t)$$

Testing the cavity predictions for the clustering transition

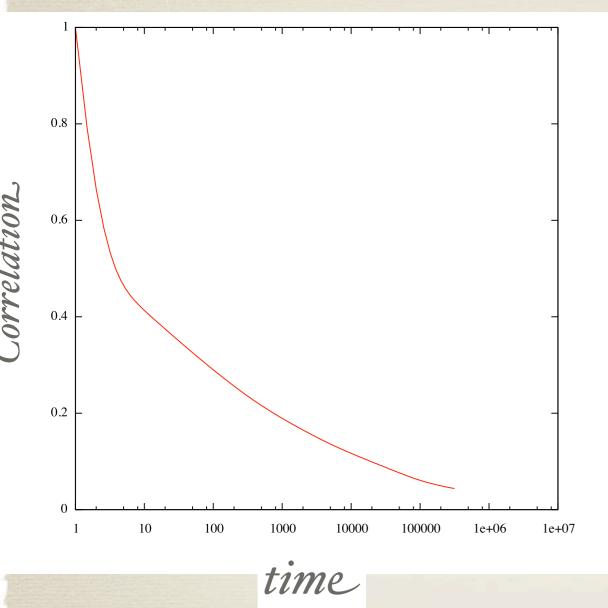


Usual Approach:

- 1) Start with a random initial condition.
- 2) compute the correlation function:

$$C(t) = \frac{1}{N} \sum_{i=1}^{N} S_i(t_{init} = 0) S_i(t)$$

Testing the cavity predictions for the clustering transition

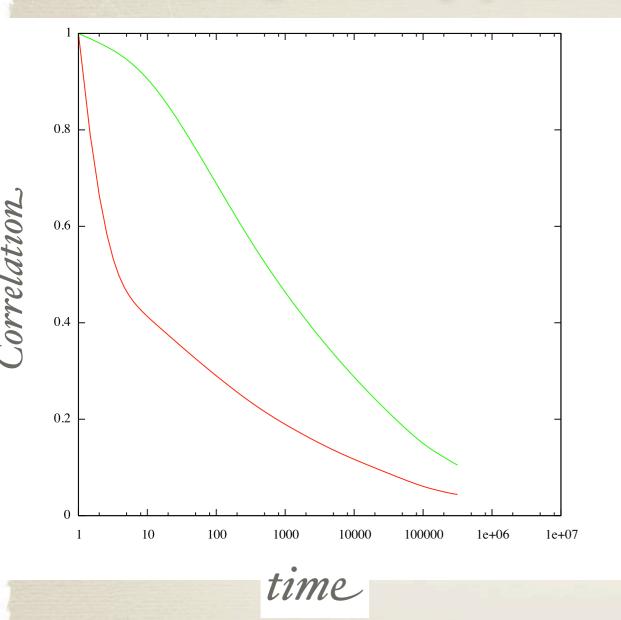


Usual Approach:

- 1) Start with a random initial condition.
- 2) Try to find an equilibrium configuration.
- 2) compute the correlation function:

$$C(t) = \frac{1}{N} \sum_{i=1}^{N} S_i(t_{init} = t_w) S_i(t - t_w)$$

Testing the cavity predictions for the clustering transition



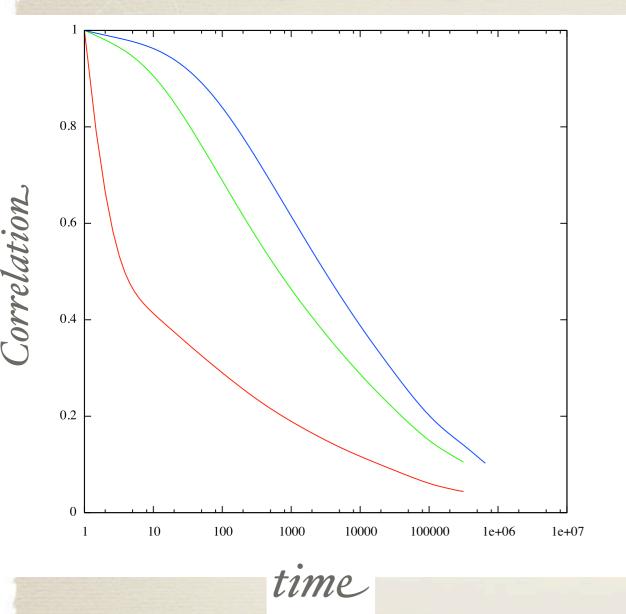
Usual Approach:

- 1) Start with a random initial condition.
- 2) Try to find an equilibrium configuration.
- 2) compute the correlation function:

$$C(t) = \frac{1}{N} \sum_{i=1}^{N} S_i(t_{init} = t_w) S_i(t - t_w)$$

tw=10

Testing the cavity predictions for the clustering transition



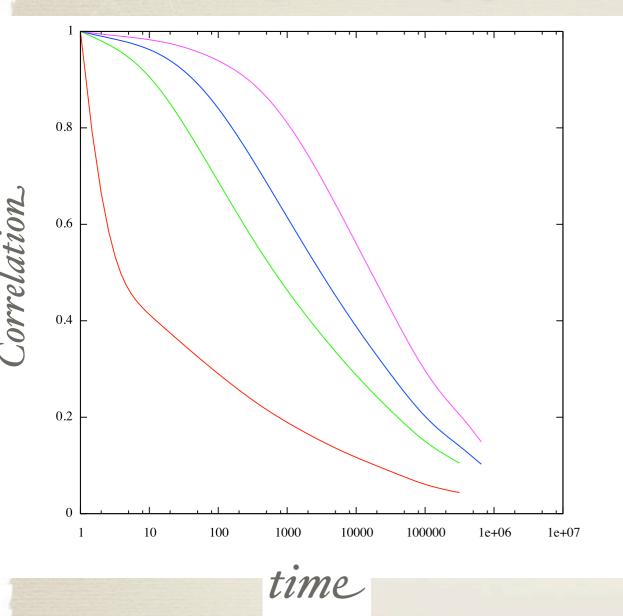
Usual Approach:

- 1) Start with a random initial condition.
- 2) Try to find an equilibrium configuration.
- 2) compute the correlation function:

$$C(t) = \frac{1}{N} \sum_{i=1}^{N} S_i(t_{init} = t_w) S_i(t - t_w)$$

tw=100

Testing the cavity predictions for the clustering transition



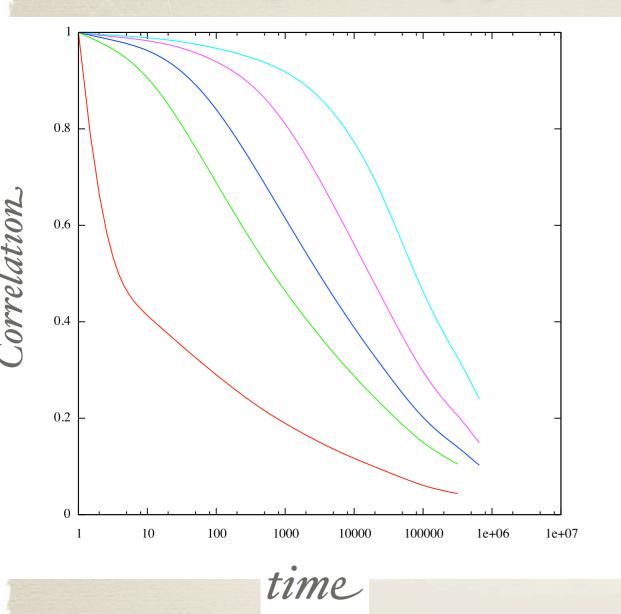
Usual Approach:

- 1) Start with a random initial condition.
- 2) Try to find an equilibrium configuration.
- 2) compute the correlation function:

$$C(t) = \frac{1}{N} \sum_{i=1}^{N} S_i(t_{init} = t_w) S_i(t - t_w)$$

tw=1000

Testing the cavity predictions for the clustering transition



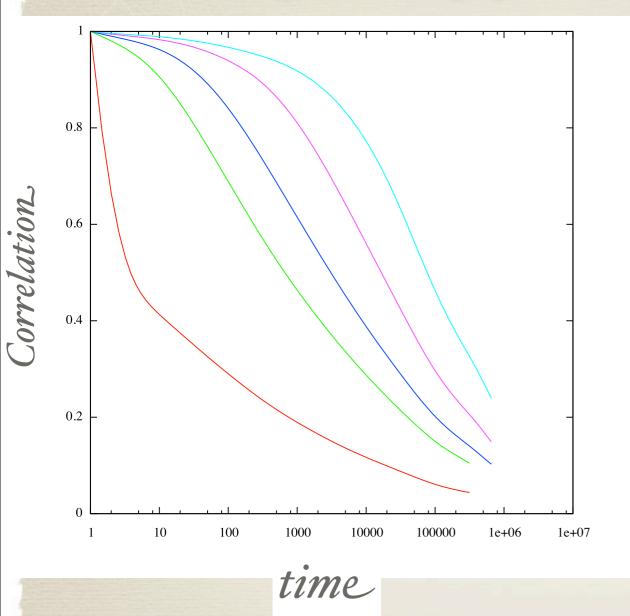
Usual Approach:

- 1) Start with a random initial condition.
- 2) Try to find an equilibrium configuration.
- 2) compute the correlation function:

$$C(t) = \frac{1}{N} \sum_{i=1}^{N} S_i(t_{init} = t_w) S_i(t - t_w)$$

tw=10000

Testing the cavity predictions for the clustering transition

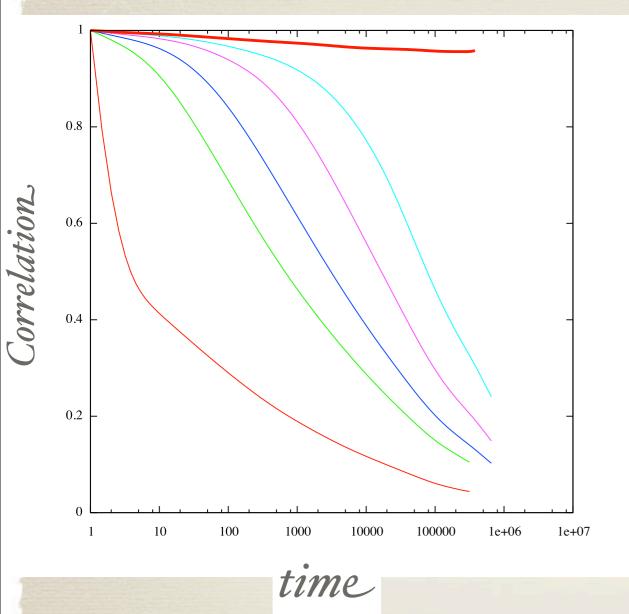


A better Approach:

- 1) Start with an equilibrated initial condition.
- 2) compute the correlation function:

$$C(t) = \frac{1}{N} \sum_{i=1}^{N} S_i(t_{init} = 0) S_i(t)$$

Testing the cavity predictions for the clustering transition

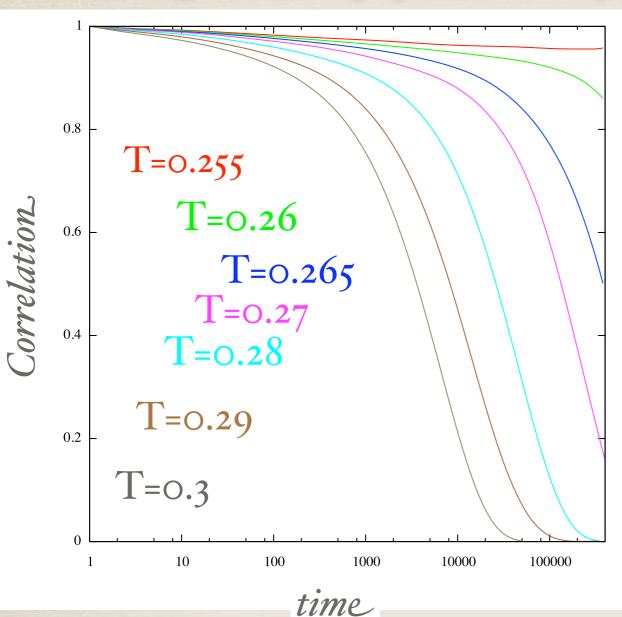


A better Approach:

- 1) Start with an equilibrated initial condition.
- 2) compute the correlation function:

$$C(t) = \frac{1}{N} \sum_{i=1}^{N} S_i(t_{init} = 0) S_i(t)$$

Testing the cavity predictions for the clustering transition



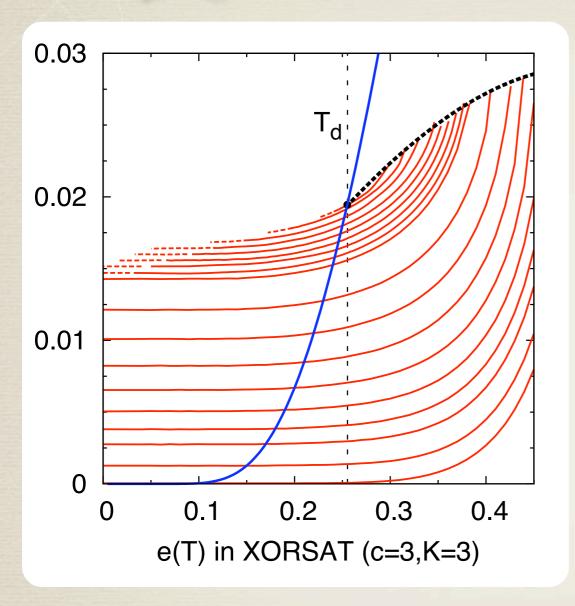
A better Approach:

Start with an <u>equilibrated</u> initial condition.

Many temperatures:

Divergence of the relaxation time

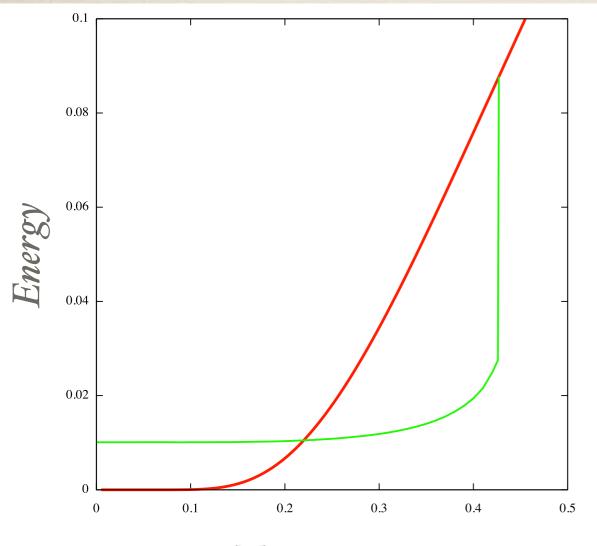
Studying Monte-Carlo annealings starting from equilibrium



XOR-SAT problems (Parity-check)

$$\mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2}$$

Studying Monte-Carlo annealings starting from equilibrium



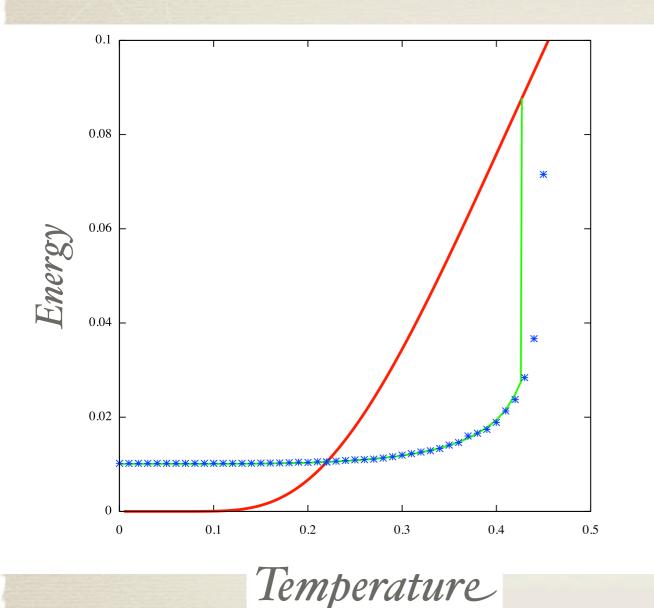
XOR-SAT problems (Parity-check)

$$\mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2}$$

N=200 000 spins

Temperature

Studying Monte-Carlo annealings starting from equilibrium

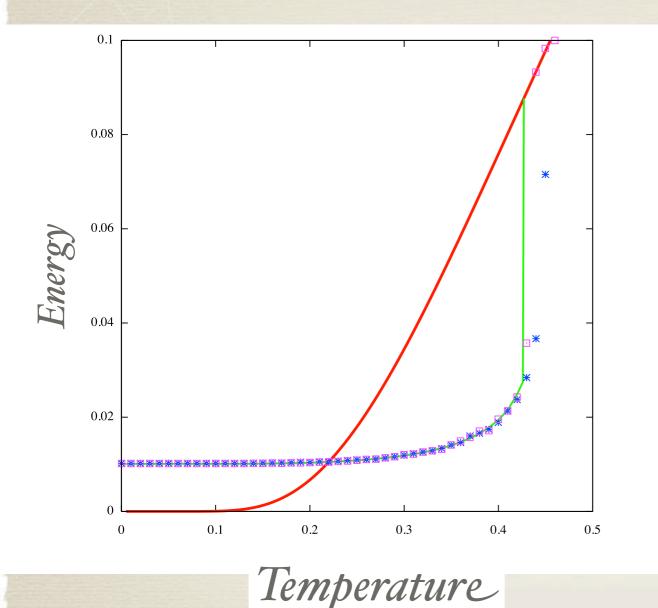


XOR-SAT problems (Parity-check)

$$\mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2}$$

N=200 000 spins

Studying Monte-Carlo annealings starting from equilibrium

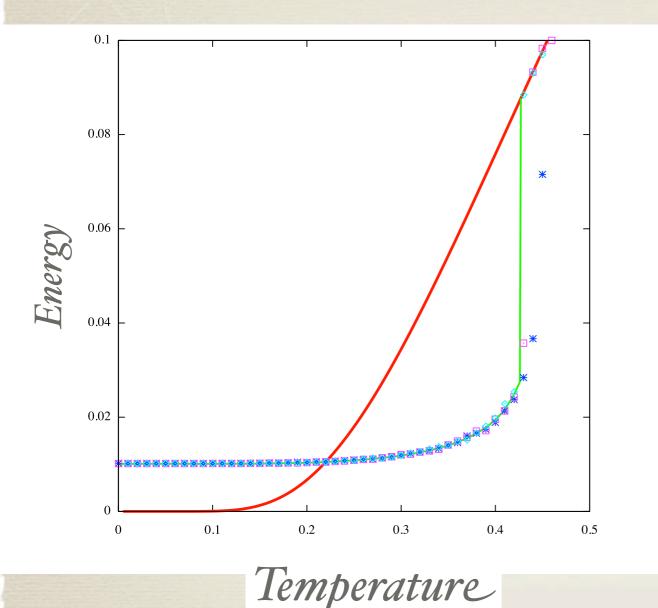


XOR-SAT problems (Parity-check)

$$\mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2}$$

N=200 000 spins

Studying Monte-Carlo annealings starting from equilibrium



XOR-SAT problems (Parity-check)

$$\mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2}$$

N=200 000 spins

Studying more complex Hamiltonians at low temperature

$$\mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2}$$

Studying more complex Hamiltonians at low temperature

$$\mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2} + \Gamma \mathcal{H}_{perturb}$$

Start with an equilibrated configuration at Γ =0 and increase Γ

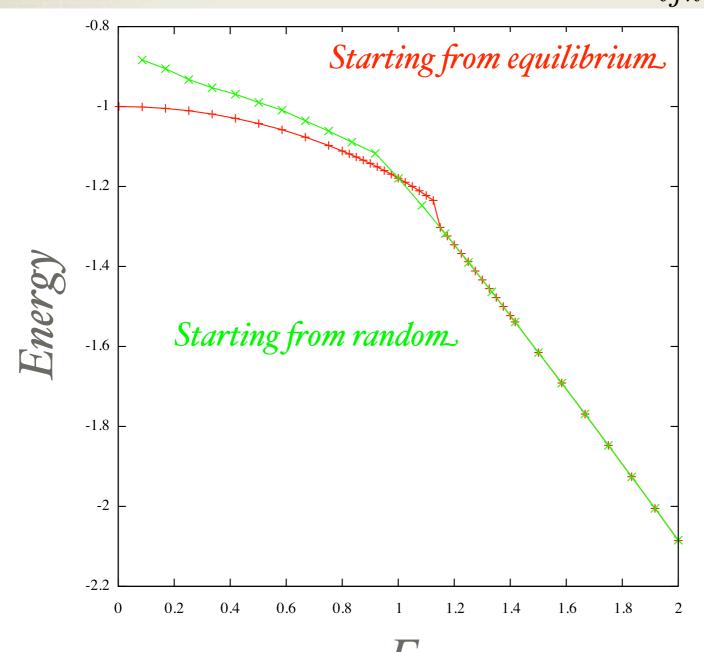
Studying more complex Hamiltonians at low temperature

Example: include a quantum transverse field!

$$\mathcal{H}(\{S\}) = \sum_{ijk} \frac{1 + J_{ijk} S_i S_j S_k}{2} \qquad \mathcal{H} = \sum_{ijk} \frac{1 + J_{ijk} s_i^z s_j^z s_k^z}{2} + \Gamma \sum_i s_i^x$$

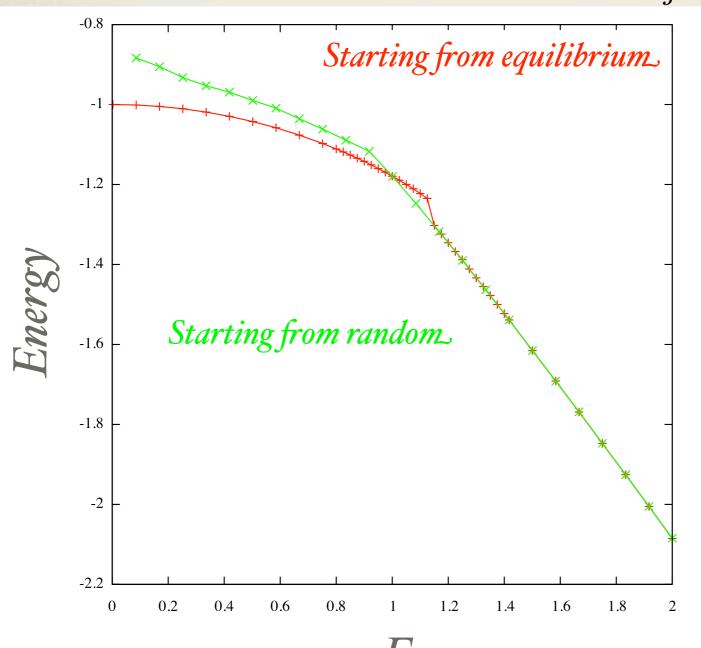
Studying more complex Hamiltonians at low temperature

$$\mathcal{H} = \sum_{ijk} \frac{1 + J_{ijk} s_i^z s_j^z s_k^z}{2} + \Gamma \sum_i s_i^x$$



Studying more complex Hamiltonians at low temperature

$$\mathcal{H} = \sum_{ijk} \frac{1 + J_{ijk} s_i^z s_j^z s_k^z}{2} + \Gamma \sum_i s_i^x$$



First order Quantum transition



Imply the failure of Quantum Annealing (or Quantum Adiabatic Algorithm)

Conclusions

A quiet planting is possible in many models.

- "Quiet" Planting does not change the properties of the ensemble up to the condensation threshold.
- Planted solutions are hard to find until the Kesten-Stigum threshold.
- Possibility to hide solutions (even a unique solution)

FK and L. Zdeborová:

- * Phys. Rev. Lett. 102, 238701 (2009)
- * arXiv:0902.4185, submitted in SIAM Journal on Discrete Mathematics
- * And more to come...

Conclusions

A quiet planting is possible in many models.

- "Quiet" Planting does not change the properties of the ensemble up to the condensation threshold.
- Planted solutions are hard to find until the Kesten-Stigum threshold.
- Possibility to hide solutions (even a <u>unique</u> solution)

There is a free lunch: instantaneous simulations.

•Many "mean field" models and random optimization models can be simulated efficiently using planting at zero or finite temperature.

FK and L. Zdeborová:

- * Phys. Rev. Lett. 102, 238701 (2009)
- * arXiv:0902.4185, submitted in SIAM Journal on Discrete Mathematics
- * And more to come...